

# Automatic Enrichments

Orfeas Menis - Mastromichalakis

# Enrichments



**Task:** enrich 20.000 records with **WIKIDATA** and **GETTY** URIs of the corresponding terms provided to us (see [this sheet](#)).

**Record fields** to be enriched: dc:description, dc:title, dc:creator, dc:contributor, dc:subject, dc:type and dcterms:spatial.

## Progress:

1. Creation of a fully multilingual thesaurus based on the provided terms and translations, enriched with Wikidata and Getty translations where needed.
2. Enrichment of descriptions almost completed. Fine-tuning of the tools and completion of enrichments by the end of February.

# Multilingual Sample Enrichments

## English Text

*Xiuzhen miyao*, a gymnastic (*daoyin/qigong*) text of unknown origin, was rediscovered and published with a preface by Wang Zai in 1513 (8th year of the Zhengde reign period of the Ming dynasty). It records 49 exercises. This illustration depicts *Tongzhi bai Guanyin* (The infant prays to the Bodhisattva Avalokitesvara), a technique used to treat pain in the heart and chest. It is practised as follows: sitting still with legs extended and hands resting on the tops of the thighs, one concentrates the attention and circulates Qi for 12 breaths.

1

- ✗ <http://www.wikidata.org/entity/Q715162>  
Guan Yin
  - ✗ <http://www.wikidata.org/entity/Q996>  
baby
  - ✗ <http://vocab.getty.edu/aat/300189561>  
infants
  - ✗ <http://www.wikidata.org/entity/Q178149>  
Bodhisattva
  - ✗ <http://vocab.getty.edu/aat/300264360>  
bodhisattas
  - ✗ <http://www.wikidata.org/entity/Q366134>  
chest
  - ✗ <http://vocab.getty.edu/aat/300039119>  
coffrets (chests)
- +

Bildtext i boken Bortom bergen: 112) Förste kejsaren - **Qin** Shi Huangdi - enade **Kina** år 221 före vår tideräkning. Hans grav bevakades av en väldig **armé** under marken.

Två **terrakotta** hästar till hälften synliga under utgrävning.

Publicerad på sid 113 i boken: Bortom bergen, Norstedts, 1983

- Småbild, 135-format
- Svartvit
- Negativ

1

- ✗ <http://www.wikidata.org/entity/Q7183>  
Qin dynasty
  - ✗ <http://www.wikidata.org/entity/Q148>  
People's Republic of China
  - ✗ <http://vocab.getty.edu/aat/300137542>  
army
  - ✗ <http://www.wikidata.org/entity/Q37726>  
army
  - ✗ <http://vocab.getty.edu/aat/300010669>  
terracotta (clay material)
  - ✗ <http://www.wikidata.org/entity/Q60424>  
terracotta
- +

## Swedish Text

# Validation

A validation interface will be integrated into our enrichment system and it will be provided to the content providers/aggregators to evaluate the produced annotations.

The validators will have the options to accept/reject an annotation, or manually create new annotations (current interface can be seen in the previous slide).

The validations will be used according to a strategy that will be agreed (e.g. reject low-confidence annotations, or reject annotations in specific language-field combinations according to evaluation) to filter out poor-quality enrichments.

As soon as the enrichments are finished and the validation interface is ready, the content providers/aggregators will be contacted. There is going to be a walkthrough of the validation procedure through our system and the validators will be guided throughout the process.



Thank you